ÖREBRO
UNIVERSITET

Lösning till tentamen i

Matematisk statistik och sannolikhetslära

MA506G

2021-08-16

---

1. En maskin som tillverkar tärningar har blivit felprogrammerad. På varje sida målas en till sex prickar med samma sannolikhet, men sidorna målas oberoende av varandra och kan därför ha flera sidor med samma antal prickar. Vi har fått en av dessa tärningar, men har inte granskat dess sidor.

   (a) Låt X vara antalet sidor som visar 6 på tärningen. Vilken fördelning har X? [3p]

   (b) Vi slår tärningen. Bestäm sannolikheten att vi slår en sexa. [3p]

   (c) Vi slår tärningen två gånger. Bestäm sannolikheten att vi slår sex i båda slagen. [4p]

   Lösning:

   (a) Eftersom sidorna är oberoende och alla sidor har samma sannolikhet att visa 6 har vi en binomialfördelning, närmare bestämt $X \sim \mathrm{Bin}(6, 1/6)$.

   (b) Det är samma sannolikhet för alla värden, så sannolikheten är $1/6$.

   (c) Enklast är förmodligen att falluppdela på antalet sexor på tärningen. Vi får då

$$\begin{aligned}\mathbf{P}(\text{två sexor}) &= \sum_{k=0}^{6} \mathbf{P}(\text{två sexor}|X=k)\mathbf{P}(X=k) \\ &= \sum_{k=0}^{6} \frac{k^2}{6^2}\mathbf{P}(X=k) \\ &= \frac{1}{36}\sum_{k=0}^{6} k^2\mathbf{P}(X=k) = \frac{\mathbf{E}(X^2)}{36}.\end{aligned}$$

   Eftersom $\mathbf{E}(X^2) = \mathbf{V}(X) + \mathbf{E}(X)^2 = 5/6 + 1 = 11/6$ får vi att sannolikheten att slå två sexor är $11/216$, vilket är betydligt högre än $6/216$ som gäller för en vanlig tärning.

2. Låt slumpvariabeln X ha fördelningsfunktionen

$$F_X(x) = \begin{cases} 1 - \frac{1}{x^2} & \text{för } x \geq 1, \\ 0 & \text{annars.} \end{cases}$$

   (a) Beräkna den betingade sannolikheten $G(x) = \mathbf{P}(X \leq 5 + x | X > 5)$. [7p]

   (b) Visa att $G(x)$ är en fördelningsfunktion genom att visa egenskaperna för fördelningsfunktioner i formelsamlingen. [3p]

Lösning:

(a) För $x > 0$ har vi

$$\begin{aligned} G(x) &= \mathbf{P}(X \leq 5 + x | X > 5) \\ &= \frac{\mathbf{P}(X \leq 5 + x \cap X > 5)}{\mathbf{P}(X > 5)} \\ &= \frac{F_X(5+x) - F_X(5)}{1 - F_X(5)} \\ &= \frac{1 - \frac{1}{(5+x)^2} - (1 - \frac{1}{5^2})}{1 - (1 - \frac{1}{5^2})} \\ &= \frac{\frac{1}{5^2} - \frac{1}{(5+x)^2}}{\frac{1}{5^2}} \\ &= 1 - \frac{25}{(5+x)^2}. \end{aligned}$$

Det följer också direkt från definitionen att $G(x) = 0$ för $x \leq 0$.

(b) Det är de tre första egenskaperna under rubriken *Fördelningsfunktion* i formelsamlingen som vi behöver visa. Genom att utgå från formlerna för $G(x)$ för negativa och positiva x får vi

$$\lim_{x \to -\infty} G(x) = \lim_{x \to -\infty} 0 = 0$$

och

$$\lim_{x \to \infty} G(x) = \lim_{x \to \infty} \left(1 - \frac{25}{(5+x)^2} \right) = 1.$$

Vidare ser vi att för $x \geq 0$ är den andra termen avtagande, vilket gör att hela uttrycket är växande. Därmed är $G(x)$ en fördelningsfunktion.

3. Weibullfördelningen är uppkallad efter Waloddi Weibull, som var professor vid KTH. En slumpvariabel $X \sim W(\lambda, \beta)$ är Weibullfördelad med parametrarna λ och β om den har fördelningsfunktionen [10p]

$$F_X(x) = \begin{cases} 1 - \mathrm{e}^{-(\lambda x)^\beta} & \text{för } x \geq 0, \\ 0 & \text{annars.} \end{cases}$$

Ett system innehåller två komponenter vars livslängder är oberoende och Weibullfördelade: $X_1, X_2 \sim W(\lambda, \beta)$. Systemet fungerar så länge båda komponenterna fungerar. Bestäm fördelningsfunktionen för systemet.

Lösning: Systemet slutar fungera så snart en av komponenterna slutar fungera. Sannolikheten att den första komponent fungerar vid tiden t är $1 - F_{X_1}(t)$, och sannolikheten att båda fungerar är $(1 - F_{X_1}(t))(1 - F_{X_2}(t))$.

Låt Y vara systemets livslängd. Vi har då

$$F_Y(t) = 1 - (1 - F_{X_1}(t))(1 - F_{X_2}(t)) = 1 - \mathrm{e}^{-2(\lambda t)^\beta},$$

vilket ger att $Y \sim W(2^{1/\beta}\lambda, \beta)$.

4. Nederbördsmängden, mätt i mm, under en dag kan ses som en stokastisk variabel $X = YZ$, där $Y \sim \mathrm{Be}(p)$ och $Z \sim \mathrm{Exp}(\lambda)$ är oberoende.

(a) Beräkna $\mathbf{E}(X)$ och $\mathbf{V}(X)$ som funktion av p och λ. [5p]

(b) Anta att nederbörden under en följd av dagar är oberoende och av samma fördelning som X ovan. Faktiska data indikerar att $\mathbf{E}(X) = 1.5$ och $\mathbf{V}(X) = \frac{27}{4}$, vilket svarar mot $p = 1/2$ och $\lambda = 1/3$. Beräkna sannolikheten att årsnederbörden överstiger 500 mm. [5p]

Lösning:

(a) Eftersom Y och Z är oberoende har vi

$$\mathbf{E}(X) = \mathbf{E}(YZ) = \mathbf{E}(Y)\underline{(}Z) = \frac{p}{\lambda}$$

och

$$\begin{aligned}
\mathbf{E}(X^2) &= \mathbf{E}((YZ)^2) = \mathbf{E}(Y^2)\mathbf{E}(Z^2) \\
&= (\mathbf{V}(Y) + \mathbf{E}(Y)^2)(\mathbf{V}(Z) + \mathbf{E}(Z)^2) \\
&= (p(1-p) + p^2)(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}) \\
&= \frac{2p}{\lambda^2},
\end{aligned}$$

vilket ger

$$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \frac{2p}{\lambda^2} - \frac{p^2}{\lambda^2} = \frac{p(2-p)}{\lambda^2}.$$

(b) Låt X_j beteckna nederbörden dag j och låt $S_k = \sum_{j \leq k} X_j$. Under antagandet att nederbörden olika dygn är oberoende och likafördelad ger centrala gränsvärdessatsen att S_{365} är ungefärligen normalfördelad med väntevärde $365 \cdot 1.5 = 547.5$ och varians $365 \cdot 6.75 = 2463.75$, eftersom vi summerar över många slumpvariabler. Vi får

$$\begin{aligned}\mathbf{P}(S_{365} > 500) &= \mathbf{P}\left(\frac{S_{365} - 547.5}{\sqrt{2463.75}} > \frac{500 - 547.5}{\sqrt{2463.75}}\right) \\ &\approx 1 - \Phi\left(\frac{500 - 547.5}{\sqrt{2463.75}}\right) \\ &= 1 - \Phi(-0.956) = \Phi(0.956) = 0.83.\end{aligned}$$

5. En tärning i form av en ikosaeder har 10 sidor, märkta från 0 till 9. Det florerar en misstanke att tärkningen är skev. Låt p vara sannolikheten att slå 9.

(a) Mikael slår tärningen 20 gånger och räknar antalet gånger som han får en nia, vilket blir 5. Gör en ML-skattning av p utifrån detta. [5p]

(b) Niklas slår tärningen tills han har fått 4 nior, vilket sker i kast 3, 9, 13 och 18. Gör en ML-skattning utifrån denna information. [5p]

Lösning:

(a) Låt X vara antalet nior Mikael får. Vi har då $X \sim \mathrm{Bin}(20, p)$. Därmed är likelihoodfunktionen

$$L(p) = \binom{20}{5} p^5 (1-p)^{15},$$

och

$$\ell(p) = \ln(L(p)) = \ln\left(\binom{20}{5}\right) + 5\ln(p) + 15\ln(1-p).$$

Då är p^* lösningen till

$$0 = \ell'(p) = \frac{5}{p} - \frac{15}{1-p},$$

det vill säga $p^* = 1/4$.

(b) Låt Y vara antalet kast som krävs för att Niklas ska få 4 nior. Vi har då $Y \sim \mathrm{NegBin}(4, p)$. Därmed är likelihoodunktionen

$$L(p) = \binom{17}{3} p^4 (1-p)^{14},$$

vilket som ovan ger ekvationen

$$0 = \ell'(p) = \frac{4}{p} - \frac{14}{1-p}$$

med lösningen $p^* = \frac{2}{9}$.

6. Quercetin ($C_{15}H_{10}O_7$, från *Quercus* (ek på latin)) är ett naturligt förekommande färgämne som kan köpas som kosttillskott, men som hittills inte kunnat visa sig vara verkansfullt mot någon sjukdom. Inte desto mindre genomförs en studie av halterna (mätt i mg/hg) i två äppelsorter. Från den första fås stickprovet $(x_1, \ldots, x_8)$ och från den andra stickprovet $(y_1, \ldots, y_{10})$, som antas vara observationer av slumpvariablerna $X_1, \ldots, X_8$ och $Y_1, \ldots, Y_{10}$ med fördelningarna $X_j \sim N(\mu_1, \sigma)$ och $Y_k \sim N(\mu_2, \sigma)$. Mätningarna ger [10p]

$$\begin{aligned} \bar{x} &= 4.4723 \\ \sum_{j=1}^{8} (x_j - \bar{x})^2 &= 0.1277 \\ \bar{y} &= 4.6266 \\ \sum_{k=1}^{10} (y_k - \bar{y})^2 &= 0.1997. \end{aligned}$$

Bestäm ett 95 % konfidensintervall för $\mu_1 - \mu_2$ och testa sedan hypotesen $H_0 : \mu_1 = \mu_2$ med felrisk 5 % och lämplig mothypotes.

Lösning: Vi väljer mothypotesen $H_1 : \mu_1 \neq \mu_2$, vilket gör att vi söker ett tvåsidigt intervall. Den sammanvägda stickprovsvariansen är

$$s_{\mathrm{p}}^2 = \frac{S_{xx} + S_{yy}}{n_1 + n_2 - 2} = \frac{0.1277 + 0.1997}{8 + 10 - 2} = 0.02046,$$

vilket med $t_{0.0025}(16) = 2.12$ ger

$$\begin{aligned} I_{\mu_1 - \mu_2} &= \left(\bar{x} - \bar{y} \pm t_{\alpha/2}(n_1 + n_2 - 2) s_{\mathrm{p}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right) \\ &= \left(4.4723 - 4.6266 \pm 2.12 \cdot 0.1430 \sqrt{\frac{9}{40}} \right) \\ &= (-0.1543 \pm 0.1438). \end{aligned}$$

Efterom $0 \notin I_{\mu_1 - \mu_2}$ avvisar vi H_0.